Predicting Potential Electronic Serials Use

Matt Jabaily
Electronic Resources and Serials Librarian
University of Colorado Colorado Springs
June 6, 2019
Outline

• Why predict future use?
• Sources of data and indicators of demand
  – Strengths and weaknesses
  – Recently published studies
• Two case studies from my library
Where I’m From

• University of Colorado Colorado Springs
• Kraemer Family Library
• 10,000 FTE
• R2 Classification
• Part of CU System
What I do

• Manage access to electronic resources (databases, online journals, and eBooks)
• Provide statistics and analysis for all electronic resources ahead of renewal decisions
  – Gather and analyze usage statistics
  – Consider role of resource in the broader collection
Why Predict Future Use?

- Identify good values for new subscriptions
- Incentivize cancelation of poor performing when not in budget crunch
- Consider opportunity cost when evaluating current subscriptions
- Become less reliant on anecdotal sources of user demand
  - Faculty requests
  - Vendor suggestions
How Can We Study This?

- Select indicator(s) of demand to study and gather data
- Purchase access to new materials
- Check use after users have a chance to access
- Look for correlations between demand indicators and use
Usefulness of Usage Stats

- Good at telling us what isn’t being used
- Less good at telling us how much something is used
- Inflation of usage stats varies by publisher
- “Garbage In, Gospel Out” (Bucknell 2012)

- To what extent does predicting COUNTER use meaningful?
- Will improvements in COUNTER 5 help?
Potential Predictive Data Sources

• Impact factor (citation metrics)
• Interlibrary loan requests
• Usage data from similar resources
• Denial/turnaway reports
• Failed link resolver requests
Impact Factor

- Impact factor measures citation of articles in a journal relative to how many articles it publishes in a given time period
- Article use and impact factor measure influence/demand/desirability of a journal’s content
- Journals with higher impact should show higher use
Impact Factor

• “Our study indicates that there is substantial correlation between citations and reported downloads, with an $R^2$ of about .75 in a simple regression.” (Wood-Doughty et al 2017)

• Wang & Mi (2019) found a moderate positive correlation between use and impact factor, as presented at last NASIG
Impact Factor

• Strengths
  – May be useful for individual titles
  – Doesn’t require examination of individual institution’s stats

• Weaknesses
  – May not hold at smaller institutions (Wood-Doughty looked at used from 10 UC libraries)
  – Libraries may already own high impact journals
Interlibrary Loan

- Number of articles requested from other libraries
- Longest and most studied indicator of unmet demand
Interlibrary Loan

• Strengths
  – Indicator of strong interest
  – Not dependent on discovery source
  – Visible, rich data (patron/article in addition to title)

• Weaknesses
  – Reliant on user action
  – Low frequency
Interlibrary Loan

- Scott (2016) estimated COUNTER to ILL request at 17:1
- 3 yr mean COUNTER use / 3 yr mean ILL = COUNTER/ILL ratio
- Results varied by publisher

Table 3. Aggregate ratio of COUNTER/ILL uses.

<table>
<thead>
<tr>
<th>Publisher</th>
<th>Titles</th>
<th>3 yr. mean of ILL</th>
<th>3 yr. mean of COUNTER use</th>
<th>COUNTER/ILL ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elsevier SD Freedom Collection</td>
<td>588</td>
<td>14.26</td>
<td>200.68</td>
<td>14.07</td>
</tr>
<tr>
<td>Springer</td>
<td>441</td>
<td>6.83</td>
<td>110.39</td>
<td>16.16</td>
</tr>
<tr>
<td>Wiley</td>
<td>622</td>
<td>6.51</td>
<td>156.40</td>
<td>24.02</td>
</tr>
<tr>
<td>Grand total</td>
<td>1651</td>
<td>9.36</td>
<td>159.88</td>
<td>17.08</td>
</tr>
</tbody>
</table>
Interlibrary Loan

• Grabowsky et al. (2019) found correlation between usage and ILL requests and modeled the relationship
Interlibrary Loan

• Barton et al. (2018) found ILL requests “only slightly positively correlated” to later full text use

• Within broad subject category, ILL requests strongly positively correlated to later full text use.
Use of Similar Resources

• Usage stats identify high use materials
  – Buy more in high use subject areas
  – Buy more on high use platforms
Use of Similar Resources

• Strengths
  – Makes logical sense
  – Safe places to put money that needs to be spent

• Weaknesses
  – Highly speculative
  – Double down or fill gaps?
Turnaway/Denial Reports

- Recorded when a user hits a paywall or is otherwise denied access to a resource
- COUNTER 4 JR2: Access Denied to Full-Text Articles by Month, Journal and Category
- COUNTER 5 TR-J2: Journal Access Denied
Turnaway/Denial Reports

• Strengths
  – Information that might otherwise be “invisible” to librarians
  – Finds users outside of library purchased databases

• Weaknesses
  – Requires some relationship with vendor/publisher (current subscriptions)
  – Misses attempts that don’t make it to publisher/vender page
  – Records data even if full text is available from another source
Turnaway/Denial Reports

- Smith (2019) found medium correlation between ILL requests and turnaways
- Results vary widely by publisher

Table 1. Results by vendor/publisher.

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>SAGE</th>
<th>Bevier</th>
<th>Springer</th>
<th>Taylor &amp; Francis</th>
<th>Wiley</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turnaways total</td>
<td>105,446</td>
<td>5,601</td>
<td>53,141</td>
<td>17,547</td>
<td>9,141</td>
<td>20,016</td>
</tr>
<tr>
<td>Turnaways average</td>
<td>15.70</td>
<td>10.00</td>
<td>26.74</td>
<td>11.23</td>
<td>6.34</td>
<td>17.15</td>
</tr>
<tr>
<td>Turnaways std. dev.</td>
<td>36.68</td>
<td>15.32</td>
<td>56.52</td>
<td>23.36</td>
<td>9.32</td>
<td>31.61</td>
</tr>
<tr>
<td>ILL requests total</td>
<td>12,062</td>
<td>1,484</td>
<td>4,930</td>
<td>1,634</td>
<td>2,377</td>
<td>1,637</td>
</tr>
<tr>
<td>ILL requests average</td>
<td>1.80</td>
<td>2.65</td>
<td>2.48</td>
<td>1.05</td>
<td>1.65</td>
<td>1.40</td>
</tr>
<tr>
<td>ILL requests std. dev.</td>
<td>4.22</td>
<td>4.59</td>
<td>5.98</td>
<td>2.46</td>
<td>3.43</td>
<td>2.61</td>
</tr>
<tr>
<td>Turnaway/ILL requests correlation</td>
<td>0.51</td>
<td>0.56</td>
<td>0.62</td>
<td>0.49</td>
<td>0.46</td>
<td>0.45</td>
</tr>
<tr>
<td>ILL requests to turnover ratio</td>
<td>11.4%</td>
<td>26.5%</td>
<td>9.3%</td>
<td>9.3%</td>
<td>26.0%</td>
<td>8.2%</td>
</tr>
</tbody>
</table>

Figure 1. Scatterplot of turnaways and ILL requests.
Turnaway/Denial Reports

- Grabowsky et al. (2019) found almost no correlation between turnaways and future use in SAGE package

| Table 1. Descriptive statistics for criterion and predictors. |
|-------------------|-----------------|-----------------|-----------------|-----------------|
| Variable          | N               | Mean            | Maximum         | Minimum         | SD              |
| usage2016         | 572             | 31.38           | 476             | 0               | 42.97           |
| Denials2015       | 572             | 0.32            | 59              | 0               | 2.89            |
| ILL_Requests2015  | 572             | 0.97            | 24              | 0               | 2.34            |

- Study found 183 turnaways (compared to 555 ILL requests)
- Very different ILL to turnaway ratio than Smith (303.3% v 26.5%)
Failed Link Resolver Requests

- Use Google Analytics to monitor link resolver pages
- Count of times link resolver fails to find full text

You are looking for:

Customer lifetime value determination and strategic implications for a cruise-ship company

Author: Berger, Paul D
ISSN: 1741-3447
Date: 09/2003
Volume: 11 Issue: 1 Page: 40-52

Sorry, this item is not available online. Please use options below to get it:

Request this item

UCCS students, staff, and faculty can get this item for free using Interlibrary Loan.
Failed Link Resolver Requests

• Presented in 2018 NASIG Snapshot Session
• Used failed link resolver requests to identify problems with holdings and possible new acquisitions
• Identified CINAHL Complete as subscription upgrade that might lead to high use
Failed Link Resolver Requests

- Knowlton, Kristanciuk, and Jabaily (2015) found between 13 and 35 percent of failed link resolver requests led to ILL requests
- Small scope and limited sample

### Table 3. ILL Requests in First Half of 2013 from Selected Titles

<table>
<thead>
<tr>
<th></th>
<th>Unique Articles Viewed</th>
<th>Total Pageviews</th>
<th>ILL Requests</th>
<th>Conversion Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Titles cancelled after 2012*</td>
<td>24</td>
<td>44</td>
<td>6</td>
<td>13.6%</td>
</tr>
<tr>
<td>B. Never subscribed</td>
<td>94</td>
<td>177</td>
<td>62</td>
<td>35.0%</td>
</tr>
<tr>
<td>TOTAL OF A &amp; B</td>
<td>118</td>
<td>221</td>
<td>68</td>
<td>30.7%</td>
</tr>
<tr>
<td>Currently subscribed</td>
<td>532</td>
<td>921</td>
<td>8</td>
<td>0.87%</td>
</tr>
</tbody>
</table>

*analysis performed only on articles with a 2013 publication date
Failed Link Resolver Requests

• Strengths
  – Records requests from all library databases that use resolver
  – Rich data (all OpenURL information)

• Weaknesses
  – Requires users to try link resolver
  – Misses access attempts from non-library resources (Google)
  – I might be the only one using it
Case study #1: CINAHL Complete

- Previous subscription to CINAHL Plus with Full Text
- Decided to upgrade after looking at failed link resolver requests
- Data from 7/17-4/18 compared to full text downloads 7/18-4/19
CINAHL Complete Predictive Statistics

- Previous subscription 22,569 full text uses from 2,152 titles
- 619 added titles (new to CINAHL)
- 274 added unique titles (new to whole collection)
- 46 ILL requests from 20 titles (all added)
- 29 ILL requests from 13 titles (unique)
- 414 failed link resolver requests from 97 titles (all added)
- 295 failed link resolver requests 65 titles (unique)
CINAHL Complete: Bad Guesses

- Based on CINAHL Average Title Use
  - 22,569 uses / 2,152 title = 10.49 uses/title
  - 619 titles * 10.49 uses/title = 6491 uses
  - 274 titles * 10.49 uses/title = 2874 uses
- Based on ILL
  - 46 requests * 17 uses/request = 782 uses
  - 29 requests * 17 uses/request = 493 uses
- Based of failed link resolver requests
  - 414 views * 5? uses/view = 2070 uses
  - 295 views * 5? uses/view = 1475 uses
CINAHL Complete Results

• 4593 full text uses from all added titles
• 1037 full text uses from unique titles
CINAHL: Uses and ILL (unique titles)

$t = 12.639$, df = 135, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval: 0.6485676 0.8045753
sample estimates: cor 0.7362028
CINAHL: Uses and Failed Link Resolver Requests

$t = 14.332, df = 135, p\text{-value} < 2.2e-16$

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval: 0.7003258 0.8356476

sample estimates: cor 0.7768016
CINAHL Complete Takeaways and Notes

• Estimates were inconsistent
• All methods showed subscription would be an acceptable value
  – <$1 - $4 a use predicted; between $1 and $2 actual
• ILL << failed link resolver requests < actual uses
• Less use of unique titles than non-unique
  – Wider date ranges for non-unique? More convenient access?
• Cannibalized use from other CINAHL titles and other full text?
Case study #2: JSTOR Collections

- Previously owned Arts & Sciences 1-8
- Acquired A&S Collections 9-11 as end of year one-time purchase
  - Some titles were previously part of another non-A&S collection
- History of high use in JSTOR
- JSTOR representatives presented evidence of turnaways
JSTOR Predictive Statistics (7/17 – 4/18)

• 40,904 JR1 downloads previous JSTOR collections
• 619 added titles from new collections
• 334 unique added titles with no other coverage
JSTOR Bad Guesses

2017-2018 JSTOR JR1 Downloads by Collection

1500-2000 downloads for each collection in A&S 9-11?
JSTOR Predictive Statistics (7/17 – 4/18)

- 44 ILL requests from added titles (31 from unique)
- 83 failed link resolver requests from added (56 from unique)
- 780 JR2 turnaways from added titles (256 from unique)
JSTOR Bad Guesses

• Based on other collections: 4,500 – 6,000 uses

• 44 ILL requests x 17 uses/request = 748 uses

• 83 failed link resolver requests x 5? uses/request = 415 uses

• 780 turnaways x 2-5? uses/turnaway = 1,560-3,900 uses
JSTOR Results

- 4907 JR1 full text requests July 2018-April 2019
  - A&S 9: 1460
  - A&S 10: 1735
  - A&S 11: 1712
- 4251 JR1 full text requests for titles added (619)
- 1439 JR1 full text requests for unique titles added (334)
JSTOR Results

2018-2019 JRSTOR JR1 Downloads by Collection

University of Colorado
Colorado Springs
JSTOR Results - ILL

\[ t = 13.386, \text{ df } = 617, \text{ p-value } < 2.2e-16 \]

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval: 0.4109600 0.5332705

sample estimates: cor 0.4744016
JSTOR Results – Failed Resolver Requests

\[ t = 13.277, \text{df} = 617, \text{p-value} < 2.2e-16 \]
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval: 0.4077266 0.5304850
sample estimates: cor 0.4713859
JSTOR Results – JR2 Turnaways

$t = 28.127$, $df = 617$, p-value $< 2.2e-16$
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval: 0.7128532 0.7821558
sample estimates: cor 0.7495513
JSTOR Takeaways

- Previous collections give us a reasonable estimate of use
- Usefulness of data sources likely depends on how users discover and access the content
- Not enough ILL and failed resolver links to know much
- Turnaways much more useful than in Grabowsky et al. (2019)
My General Takeaways

• Ample evidence of unmet demand is useful in predicting potential use
• Scant evidence of demand is less useful
• Reliance on ILL data alone may be too conservative
• Examining data for clear values may be fruitful
Next steps

• Revisit data with greater statistical rigor
• Look at interactions of multiple indicators
• Look at other new acquisitions for more clarity
• Look for new/better data from COUNTER 5
Questions?

• mjabaily@uccs.edu
References


References


Wood-Doughty, A., Bergstrom, T., & Steigerwald, D. (2017). *Do download reports reliably measure journal usage? Trusting the fox to count your Hens?* Retrieved from [https://escholarship.org/uc/item/1f221007](https://escholarship.org/uc/item/1f221007)